8 Data Quality

The quality of the collected data has been carefully assessed. The data are individually recorded in the software model for each process description, and can be described in terms of their precision, completeness, accurate representation of physical systems, consistency, and quality of their source or origin.

Precision Different levels of precision were achieved, depending on the origin of the data. The soybean agriculture model contains measured data, with a high level of precision. The emissions data for the biodiesel also come from measured data with a high level of precision. The remaining data consist of secondary and estimated data, with a poor level of precision (see Origin of the data below).

Completeness The completeness, as defined by International Standards Organization (ISO)[93], is very good for the soybean agriculture modeling. Also, the petroleum refining model contains averaged data for all U.S. refining operations. The data describing soybean crushing and conversion to biodiesel are based on data describing specific facilities and may not represent the average data of a diversified industry.

Representativeness No attempt has been made to precisely estimate the representativeness of the systems under study. However, the representativeness for the agriculture modeling is good except that it is for 1990 data. The soybean crushing data are thought to represent a typical crushing facility. The same is true for crude oil extraction (although the data are from 1983). Soybean oil conversion is felt to represent an average facility, as are the crude oil refining data.

Consistency The data for individual processes have been collected by different members of the project team. All data, however, have been captured into the software by one person on the project team who went over all these data individually, double checking with other persons of the team as needed. These procedures help ensure a certain degree of consistency among collected data. However, each source of data is different, and some interpretation had to be done.

Origin of the data The soybean agriculture and petroleum refining data come from actual measured data that were aggregated and presented in secondary sources. Also, the emissions for biodiesel use are from engine testing measured data. The soybean crushing and soybean oil conversion data are a mix of measured and modeled data. The remaining data are from secondary sources and modeled data. Nevertheless, whenever actual data were available, they have been preferred to modeled data. In any case, the origin of the data, including whether the data are measured, calculated, or estimated, is recorded in the software model.

[93] ISO/DIS 14041: 1997(E), Version 05/03/97.

This page left intentionally blank